学 位 論 文 の 要 旨

氏名　　岩田　春子

学　位　論　文　名	Maintenance and Discovery of Domain Knowledge for Nursing Care using Data in Hospital Information System
発　表　雑　誌　名 （巻，初頁～終頁，年）	Fundamenta Informaticae (137, 237-252, 2015)
著　　者　　名	Haruko Iwata, Shoji Hirano and Shusaku Tsumoto

論 文 内 容 の 要 旨

INTRODUCTION

Twenty years have passed since clinical data were stored electronically as a hospital information system(HIS). Stored data give all the histories of clinical activities in a hospital, including accounting information, laboratory data and electronic patient records. Due to the traceability of all the information, a hospital cannot function without the information system. All the clinical inputs are shared through the network service in which medical staff can retrieve their information from their terminals. Since all the clinical data are distributed stored and connected as a large-scale network, HIS can be viewed as a cyberspace in a hospital: all the results of clinical actions are stored as "histories". It is expected that similar techniques in data mining, web mining or network analysis can be applied to the data. Dealing with cyberspace in a hospital will give a new challenging problem in hospital management in which spatiotemporal data mining, social network analysis and other new data mining methods may play central roles.

This paper proposes a temporal data mining method to maintain a clinical pathway used for schedule management of clinical care. Since the log data of clinical actions and plans are stored in hospital information system, these histories give temporal and procedural information about treatment for each patient. The method consists of the following four steps: first, histories of nursing orders are extracted from hospital information system. Second, orders are classified into several groups by using clustering and multidimensional scaling method. Third, by using the information on groups, feature selection is applied to the data and important features for classification are extracted. Finally, original temporal data are split into several groups and the first step will be repeated. The method was applied to a dataset whose patients had an operation of cataracta. The results show that the reuse of stored data will give a powerful tool for

maintenance of clinical pathway, which can be viewed as data-oriented management of nursing schedule.

METHODS

The proposed method is composed of the following three processes: data preparation, data mining and clinical pathway construction. Data is prepared by two-staged datawarehousing: first, target dataset are extracted from hospital information system into conventional datawarehouse, which is called first DWH. Then, the data is split into two DWHs: contents DWH and histories DWH. Temporal datasets of the number of orders will be stored in the secondary DWH. Data mining process is applied to the generated data sets in this secondary DWH. It is notable that the DWH process proposed here is domain independent, except for parameter settings. Secondly, temporal data mining process is composed of the following three steps. Since data from secondary DWH can be viewed as numerical temporal sequences, in which samples consist of nursing orders and attribute consists of each date, similarities among sequences are calculated. Then, clustering and multidimensional scaling (MDS) are applied and grouping of nursing orders is obtained. Next, obtained labeled is used for rule induction, where selects the features important for classification of the labels. Then, by using features, an original table is decomposed into small tables. If the grouping is the same as the original one, the algorithm quits. Otherwise, clustering, MDS and rule induction are applied in a recursive way. Finally, by using evaluation criteria, nursing orders used in decomposed clusters are compiled into a clinical pathway.

For experimental evaluation, the Ward method for clustering, the metric classical Torgerson method for MDS and C4.5 for rule induction implemented R 2.15.1 were used.

RESULTS AND DISCUSSION

Experimental evaluations are two-fold. In the first experiments, we focus on histories of patients who were admitted to the hospital for operations of cataracta. The number of patients is 121 in 2010. For this disease, the clinical pathway mentioned in the above section is used to optimize the schedule of treatment. The nursing orders are counted during the stay of each patient and regard chronological change of each order as a temporal sequence. In the first step, clustering results give two major groups: one includes the orders indispensable to this disease and the other includes those which are rather specific to the status of each patient, except for preoperation instruction. MDS gives further classification of the first group into three subgroups and the second one into two subgroups The former three subgroups consist of vital signs (BP, BT and PR), body care (Coaching and Wash), and watchlist (Eye Symptoms and Nausea). The latter two groups consist of preoperation instruction and other symptoms which may specific to the

status to the patients. The results also give another interesting observation: comparison of the above results with the existing clinical pathway shows that the pathway lacks an order, Wash, which is similar to Coaching, because their temporal patterns are very similar to other orders indispensable to the treatment. So, they should be included into clinical pathway.

With the labels obtained, rule induction methods can be applied to the dataset. The obtained features show that for d1, d2 and d3 give complete classification for the labels, and preoperation, postoperation, d4 and d5 do not have enough information for complete classification. Thus, the data should be split into three groups: the first interval is preoperationand operation, the second one is d1 to d3, and the final interval is d4 and d5. In the next cycle, the same procedures will applied to each subset of data. Then, clustering is applied to each decomposed data table in a recursive way.

Then, the proposed method was evaluated on data on nursing orders extracted from hospital information system, which were collected from Apr.1, 2009 to Mar. 31, 2010. The target diseases were selected from 10 frequent diseases whose patients were admitted to the university hospital during this period and where a corresponding given clinical pathway was applied. The results obtained are shown in the following three tables. The results show that the method is able to construct a clinical pathway for each disease. Furthermore, the best three major intervals suggested the optimal and maximum length of stay, although information on frequency of nursing orders needed to determine the optimal length of stay. For example, in the case of cataracta, the length of stay estimated from three major intervals is 5 days, but with frequency information, the fourth date has smaller frequency of orders, compared with other intervals. Thus, the optimal length will be estimated as 4 days (0,1,2,3).

CONCLUSIONS

In this paper, we propose a general framework on innovation of hospital services based on temporal data mining process. This process can be called similarity-based visualization approach in which similarity based methods, such as clustering and multidimensional scaling (MDS) and correspondence analysis. We applied the process to datasets of #nursing orders for cases for operation of cataracta where clinical pathway has been introduced. By using Clustering and MDS, we obtained two major groups in the nursing orders: ones were indispensable to the treatment, and the others were specific to the status of patients. Then, in the step for feature selection, the first day of postoperation could be viewed as a threshold in the original datasets. Thus, periods before and after operation should be dealt as independent datasets. Repeating these steps, we could characterize the temporal aspects of nursing orders, and then found missing information in the existing pathway.

論文審査及び最終試験又は学力確認の結果の要旨

㊙ ・ 乙	氏　名	岩田 春子
学位論文名	Maintenance and Discovery of Domain Knowledge for Nursing Care using Data in Hospital Information System	
学位論文審査委員	主　査	木下　芳一　印
	副　査	礒部　威　印
	副　査	大平　明弘　印

論文審査の結果の要旨

医療の標準化と質保証のためにクリニカルパスは重要であるが、その作成は医師と看護師の手作業によるもので、作成後，実症例による最終的なカスタマイズが必要であった。申請者らは、クリニカルパス作成の支援を行うため、電子化された診療歴の解析によるクリニカルパスの自動作成を行う方法を提案した。本方法では、オーダ実行数に関する時系列データを島根大学医学部附属病院の病院情報システムから抽出し、多次元尺度構成法およびクラスタ分析を適用し、診療に必須のオーダの選定を行っている。さらに選定された方法から、逆に選定の根拠となる病日それぞれの貢献度を算出している。貢献度によって、病日毎にデータ分割し、それぞれの分割したデータに再帰的にクラスタ分析を適用している。分割したデータ毎に得られた必須オーダを時系列的に並べることでクリニカルパスの候補が生成されている。申請者らは本方法を DPC コード 020110xx97x0x1 (白内障, 水晶体の疾患　手術あり　手術・処置等 2 なし　重症度等　両眼) が付された 134 症例のうち左から右の順に白内障の手術が行われ、クリニカルパス適用があり、オーダ実施が完了した 65 症例のオーダ歴を用いて検証している。本研究により自動作成されたクリニカルパスは現場で実行中のクリニカルパスの中には記載されていないが現実には実施されているオーダを含み、対象疾患のクリニカルパスの改訂を行うことが可能となった。病院情報システムを利用してクリニカルパスの生成を支援する方法の報告はこれまでなく、本自動作成システムの実用化がクリニカルパスの生産性を高めうる重要な研究であると評価し、本研究内容は博士の学位授与に値すると判断した。

最終試験又は学力の確認の結果の要旨

申請者は病院情報システム内の診療情報を用いて多元尺度構成法とクラスタ分析を行ってクリニカルパスの自動作成を行う方法を開発し、その有用性を明らかとした。関連領域の知識も豊富であり、博士の学位に値すると判定した。　(主査　木下芳一)

申請者は電子カルテ内の看護オーダ履歴を用い、クラスタ分析と多次元尺度構成によりパスの最適化を指向した自動作成を行った。臨床現場への応用が期待される研究であり、関連知識も豊富で質疑応答も的確なため、学位授与に値するものと判断した。　(副査　礒部　威)

申請者は病院情報システムから情報抽出し，クリニカルパスを自動作成する方法をクラスタ分析と多元尺度構成法により開発した。関連知識も豊富で学位授与に値するものと判断した。　(副査　大平明弘)

(備考) 要旨は、それぞれ 400 字程度とする。